Expert Declaration
in re *Free Speech Coalition v. Holder* 2:09–4607 E.D. Pa.

Philip B. Stark

4 April 2013

1   Qualifications

I am Professor and Chair of the Department of Statistics and a faculty member in the Graduate Program in Computational Science and Engineering at the University of California, Berkeley. I have published more than one hundred articles and books and have served on the editorial board of several scientific journals. I have lectured at universities and professional societies in seventeen countries. I was a Presidential Young Investigator, a Miller Research Professor, and the recipient of the Chancellor's Award for Research in the Public Interest. I am an Accredited Professional Statistician and a Certified Physicist. I have consulted for the U.S. Department of Justice, the U.S. Federal Trade Commission, the U.S. Department of Agriculture, the U.S. Census Bureau, the U.S. Department of Housing and Urban Development, the U.S. Department of Veterans Affairs, the California Secretary of State, the California Attorney General, the California Highway Patrol, the Colorado Secretary of State, and the Illinois State Attorney. I have testified to the U.S. House of Representatives Subcommittee on the Census; the State of California Senate Committee on Elections, Reapportionment and Constitutional Amendments; the State of California Assembly Committee on Elections and Redistricting; and the State of California Senate Committee on Natural Resources. I have been an expert witness or non-testifying expert in a variety of state and federal cases. I have been qualified as an expert in state and federal courts. I have testified for Plaintiffs and for De-

fendants. Appendix I is my current CV, which lists recent deposition and trial testimony.

2 Assignment

I was asked by the U.S. Department of Justice to review and comment on the expert reports of Dr. Michelle Drouin and Dr. Marc A. Zimmerman, focusing on their conclusions about the percentage of young adults in the U.S. who have sent or received "sexts." I charge $1200 per hour.

3 Materials Relied Upon

I read the two expert reports and most of the articles they cite; see below for detail.

4 Summary of Findings

Neither Dr. Drouin nor Dr. Zimmerman has a scientific or statistical basis for estimating the percentage or number of U.S. young adults who send or receive "sexts." The numbers they give are guesses, not scientific estimates.

The studies they cite attempt to measure the prevalence of sexting in groups that are (or that may be) quite unrepresentative of U.S. young adults as a whole. Moreover, their estimates have uncertainties that cannot be quantified usefully. (The true percentage is certainly between 0% and 100%, but that is hardly helpful.)

Dr. Drouin's and Dr. Zimmerman's estimates of the percentages and numbers of young adults who have sent or who have received "sexts" are derived using methods that are statistically unsound. These estimates are extrapolations from "samples of convenience" to much larger populations that may be quite unlike their samples. Decades of statistical theory and experience show that such estimates are unreliable. This is true even if some characteristics of the sample (e.g., demographics or racial characteristics) match the characteristics of the population, or if some results of the study match other studies. To extrapolate reliably, and to quantify the uncertainty usefully, it is necessary to start with a random "probability" sample from the population of interest.

If the studies on which Drs. Drouin and Zimmerman rely had sampled at random from U.S. young adults, and if the response rates in those studies had been high enough, it would have been possible to estimate these percentages and numbers in a way that is nearly unbiased, and to quantify the uncertainty in those estimates. From the data available, neither is possible.

The data and studies that Drs. Drouin and Zimmerman rely on, taken at face value, show that *some* young adults send and receive sexts, but they do not support any reliable quantitative estimate of the national percentage or national number of young adults who have sent or who have received sexts.

5   Additional detail

The following sections give the statistical basis for my conclusions. The fundamental problem with the estimates given by Drs. Drouin and Zimmerman is that their surveys did not use random sampling, much less random samples from the relevant population—U.S. young adults. In addition, the response rates to their surveys are unknown. As a result, their estimates are unreliable, and the uncertainty in their extrapolation from their samples to U.S. young adults is unknown.

In the case of Dr. Drouin's studies, there are good reasons that the sample would differ from other students at other universities, from students at other U.S. universities, and from U.S. young adults as a group. In the case of Dr. Zimmerman's studies, there are good reasons that his sample would differ from U.S. young adults as a group.

5.1   Random sampling

It is possible to draw reliable conclusions about a population from a relatively small sample of the population, if the sample is drawn appropriately. For instance, it is possible to tell whether a batch of soup is too salty by tasting only a teaspoon of soup, whether the batch is a 1-quart saucepan or a 10-gallon pot. It is possible to tell whether the bath is too hot by sticking your toe in the water.

The key is to draw the sample at random. This amounts to stirring the soup before taking the teaspoon, or to mixing the bathwater before putting your toe in. If the soup has not been stirred adequately, the teaspoon could easily be saltier than average for the pot as a whole, or less salty than the

3

pot as a whole. If the bathwater has not been mixed, your toe could hit a spot substantially warmer or colder than average. Stirring before taking the teaspoon or putting your toe in the water–i.e., sampling at random—is key.

Random samples are also called "probability samples." In a probability sample from a population of items, every item has a known chance of being selected.[1]

For probability samples, it is often possible to quantify the uncertainty in estimating a property of the population from the sample from which the sample was drawn. For instance, it is possible to estimate the likely difference between the percentage of a probability sample that has some characteristic, and the percentage of the population that has that characteristic. An example of a characteristic is "has sent or received a sext." With estimates based on probability samples, it is often possible to derive a "margin of error" for the estimate, or a confidence interval for the quantity estimated. For estimates based on other types of samples, it is generally impossible to quantify the uncertainty usefully. The notions of "margin of error" and "confidence interval" do not apply to estimates based on non-probability samples.

In contrast to a random or probability sample, a "sample of convenience" consists of items from the population that are easily accessible. In a sample of convenience, not every item from the population has an equal chance of being selected—in general, the chances are unknown, or not even defined. It is not possible to estimate the likely difference between the average of a sample of convenience and the average of the population from which that sample was drawn. It is not possible to calculate a margin of error or a confidence interval from a sample of convenience.

The studies co-authored by Dr. Drouin and by Dr. Zimmerman, and most of the studies they cite, are based on samples of convenience rather than random samples: The samples in their studies consist of students in a particular class who agreed to participate (in some cases, for course credit), or of young adults who—invited by friends of friends who had responded to a Facebook ad—signed up on the promise of $20. Because these are samples of convenience (and, in the case of Dr. Zimmerman's estimates, because the response rate is unknown and unknowable) it impossible to quantify the likely difference between the rate of sexting in their samples and the rate of

---

[1]A "simple random sample" is a special case of a probability sample. In a simple random sample of n items from a population containing N items, every subset consisting of n of the N items is equally likely to be selected. This implies that every item in the population is equally likely to be in the sample.

sexting for U.S. young adults as a group. It is not possible to calculate a margin of error or a confidence interval based on those samples. It is not possible to determine the size of the statistical bias in their estimates. These estimates are not reliable and there is no way to quantify the uncertainty of these estimates usefully.

The following quotations make it clear that probability (i.e., random) samples are the *sine qua non* for extrapolation, and that samples of convenience do not support generalizations of the kind Drs. Drouin and Zimmerman have made:

> The use of [random] sampling techniques maximizes both the representativeness of the survey results and the ability to assess the accuracy of estimates obtained from the survey. . . . [Random] sampling offers two important advantages over other types of sampling. First, the sample can provide an unbiased estimate of the responses of all persons in the population from which the sample was drawn . . . Second, the researcher can calculate a confidence interval that describes explicitly how reliable the sample estimate of the population is likely to be. . . . [Q]uantitative values computed from [samples of convenience] . . . should be viewed as rough indicators rather than as precise quantitative estimates. Confidence intervals should not be computed." (Diamond, S.S., 2000. *Reference Guide on Survey Research*, in *Federal Judicial Center Reference Manual on Scientific Evidence, 2nd edition*, Federal Judicial Center, at 242–244)

> [A]bout the only way of examining how good [a sample of convenience] may be is to find a situation in which the results are known, either for the whole population or for a [random] sample, and make comparisons. Even if a method appears to do well in one such comparison, this does not guarantee that it will do well under different circumstances." (Cochran, W.G., 2002. *Sampling Techniques*, Third Edition, John Wiley and Sons, New York, at 10)

> [Samples of convenience have] a heavy dependence on the validity of broad assumptions about the distributions of the survey variables in the population. On the contrary, from the results

> of ideal [random] sampling, the inferences to the population can
> be made entirely by statistical methods, without assumptions re-
> garding the population distributions." (Kish, L., 1965. *Survey
> Sampling*, John Wiley and Sons, New York, at 19)

Moreover, even if the sample is drawn at random, if some of the subjects
selected for the survey cannot be reached or do not submit to be surveyed,
that "nonresponse" is potentially a source of bias in the results:

> One suggested formula for quantifying a tolerable level of non-
> response in a probability sample is based on the guidelines for
> statistical surveys issued by the former U.S. Office of Statisti-
> cal Standards. According to these guidelines, response rates of
> 90% or more are reliable and generally can be treated as random
> samples of the overall population. Response rates between 75%
> and 90% usually yield reliable results, but the researcher should
> conduct some check on the representativeness of the sample. Po-
> tential bias should receive greater scrutiny when the response
> rate drops below 75%. If the response rate drops below 50%, the
> survey should be regarded with significant caution as a basis for
> precise quantitative statements about the population from which
> the sample was drawn. (Diamond, S.S., 2000. *Reference Guide
> on Survey Research*, in *Federal Judicial Center Reference Man-
> ual on Scientific Evidence, 2nd edition*, Federal Judicial Center,
> at 245)

As discussed below, the response rates in the surveys that Drs. Drouin and
Zimmerman rely on are, for the most part, either low or entirely unknown.
Dr. Drouin's papers do not report the response rate. Dr. Zimmerman's work
does not even allow the response rate to be ascertained.

5.2   The underlying studies

This section briefly reviews the studies cited by Drs. Drouin and Zimmerman,
several of which are co-authored by them (separately). Some of these studies
are not peer-reviewed academic studies; those are to be viewed with particu-
lar skepticism. As discussed below, all but five of the studies use "samples of
convenience" and have unknown response rates. There is no way to quantify

6

usefully the uncertainty in extrapolating from such samples to larger populations, such as all U.S. young adults. Notably, the studies that use probability samples report far lower rates of sexting than the studies based on samples of convenience. However, those are studies of adolescents, not U.S. young adults, and some have quite low response rates. Thus, none of the studies makes it possible to estimate the rate of sexting among U.S. young adults reliably, and to quantify the uncertainty in the estimate usefully.

Of the peer-reviewed studies, several are surveys of volunteers in undergraduate psychology classes at a single university. Such studies are not a sound basis for extrapolation.[2]

Extrapolating from students in a particular psychology course at a particular midwestern university to other students and other universities might be reasonable if the question were one of biology—where differences across individuals may not have much to do with their interests, social environment, and socioeconomic background. But it is not reasonable for issues of attitudes and behaviors.

I have been at various universities in various capacities for more than 36 years. I have taken courses at seven universities, in a wide variety of subjects. I have lectured at dozens of universities in a range of departments. I have taught "service" courses in statistics for psychology and social science majors, business majors, and engineering majors. I was undergraduate major advisor for many years, and have had one-on-one interactions with thousands of students. Firsthand experience tells me that, on the whole, psychology majors are not exactly like philosophy majors, physics majors, math majors, economics majors, English majors, art majors, music majors, or computer science majors in their attitudes and social interactions. Students at MIT are not exactly like students at Princeton, the University of Texas, UCSD, or Berkeley in their attitudes and social interactions. Students at large public universities are not exactly like students at small liberal arts colleges in their attitudes and social interactions. There is no sound basis for extrapolating from a sample of convenience of students in a psychology course at a particular midwestern university to all university students in the U.S., and certainly not to all U.S. young adults. There are good reasons to doubt the accuracy

---

[2]See, for instance, C. Wang, 1993, *Sense and Nonsense of Statistical Inference: Controversy, Misuse, and Subtlety.* Marcel Dekker, Inc., NY, at 180. Wang notes that "most findings in psychology journals (and other soft sciences as well) have to be looked at with caution," because (citing J. Sullivan as reported in S. Hite, 1987), "many if not most articles in psychology journals are based on data from college students, and then generalized."

of such extrapolation.

The Federal Judicial Center Reference Manual on Scientific Evidence offers additional relevant warnings:

> One indication that a survey offers probative evidence is that it was designed to collect information relevant to the legal controversy (e.g., to estimate damages in an antitrust suit or to assess consumer confusion in a trademark case). Surveys not conducted specifically in preparation for, or in response to, litigation may provide important information, but they frequently ask irrelevant questions or select inappropriate samples of respondents for study. Nonetheless, surveys do not always achieve their stated goals. Thus, the content and execution of a survey must be scrutinized even if the survey was designed to provide relevant data on the issue before the court. (Diamond, S.S., 2000. *Reference Guide on Survey Research*, in *Federal Judicial Center Reference Manual on Scientific Evidence, 2nd edition*, Federal Judicial Center, at 236–7)

> One of the first steps in designing a survey or in deciding whether an existing survey is relevant is to identify the target population (or universe). The target population consists of all elements (i.e., objects, individuals, or other social units) whose characteristics or perceptions the survey is intended to represent. (Ibid., at 239)

> The target population consists of all the individuals or units that the researcher would like to study. The sampling frame is the source (or sources) from which the sample actually is drawn. . . . A survey that provides information about a wholly irrelevant universe of respondents is itself irrelevant. . . . More commonly, however, the sampling frame is either underinclusive or overinclusive relative to the target population. If it is underinclusive, the surveys value depends on the extent to which the excluded population is likely to react differently from the included population. (Ibid., at 240–1)

Below is a summary of the studies Drs. Drouin and Zimmerman cite, including their own. The summary shows that none of the studies was designed

8

to answer the question at issue in this litigation; most rely on samples of convenience; most have very low or unknown response rate; and none draws a probability sample from the relevant target population—all U.S. young adults. As a result, none provides a basis for a reliable estimate of the prevalence of sexting among U.S. young adults.

1. Associated Press & MTV, 2009. AP-MTV Digital abuse study, Executive Summary.
   `http://www.athinline.org/MTVAP_Digital_Abuse_Study_Executive_Summary.pdf` Last visited 24 March 2013

   This is not a peer-reviewed study. It seems to involve a 1,244 responses from panel of volunteers who were originally recruited through random digit dialing, but the response rate in the panel recruitment phase is not reported. Whether the original panel resembles a (fair) probability sample of U.S. young adults, or of any other group, is unknown. Moreover, the response rate to this survey within the panel does not seem to have been reported.

2. Bauermeister, J. A., M.A. Zimmerman, M.M. Johns, P. Glowacki, S. Stoddard, and E. Volz, 2012. Innovative recruitment using online networks: Lessons learned from an online study of alcohol and other drug use utilizing a web-based, Respondent-Driven Sampling (we-bRDS) strategy. *Journal of Studies on Alcohol and Drugs, 73(5)*, 834–838. PMCID: PMC3410951

   This study does not use probability sampling. The participants were recruited online: a sample of convenience. With respondent-driven sampling, it is impossible to know how many people were solicited to participate, so the response rate is unknown and unknowable. Moreover, it is impossible to know how well or poorly the sampling frame matches the target population.

3. Benotsch, E.G., D.J. Snipes, A.M. Martin, and S.S. Bull, 2012. Sexting, Substance Use, and Sexual Risk Behavior in Young Adults. *Journal of Adolescent Health.* doi: 10.1016/j.jadohealth.2012.06.011

   The sample consists of 763 undergraduate volunteers in psychology classes at a large mid-Atlantic university; they received course credit for participating in the study. It is not a probability sample. The response rate was approximately 52%. Over 65% of the sample were

freshmen (generally on the low end of the 18–24 year old range); there were almost twice as many women in the sample as men; and less than 53% of the sample was white: the sample is facially unrepresentative of U.S. young adults with respect to age, gender, and race. There is no reason to believe it is representative with respect to sexting.

4. Dake, J.A., J.H. Price, L. Maziarz, and B. Ward, 2012. Prevalence and Correlates of Sexting Behavior in Adolescents. *American Journal of Sexuality Education, 7:1*, 1–15.

   This study was based on a stratified random sample of classrooms from 35 schools in three midwestern states. The opt-out rate by parents was about 4%; among students whose parents permitted them to take the survey, the response rate was about 95%, making the overall response rate (for students) approximately 90%, a total of 1,290 students. This is by far the best study among those cited, and it is rather more recent than some of the others. Notably, the percentage of students who report sexting is 17% in this sample.

   The combination of random sampling and high response rate *could* make this study reliable for extrapolating from those classrooms to the rest of the students in those 35 schools. But since the schools themselves do not appear to have been selected at random, and since they were purposely selected from only three midwestern states, the sampling frame does not match the target population for this litigation. Moreover, the target population for this study consists of adolescents, not young adults.[3] The study does not provide a reliable basis for extrapolating to U.S. young adults as a whole.

5. Drouin, M. and C. Landgraff, 2012. Texting, sexting, and attachment in college students' romantic relationships, *Computers in Human Behavior, 28*, 444–449.

   The sample consists of 744 undergraduate volunteers from a psychology course who were in a "committed relationship." It is a sample of

---

[3]The statistical tests (*p*-values, confidence intervals, etc.) reported in the study appear to have been computed incorrectly: from what I can see in the document, it seems that the authors treated their sample as if it were a *simple random sample* of students, while in fact it was a *stratified random cluster sample* of students, since the "units" selected at random were classrooms, not individual students. Hence, the confidence intervals and *p*-values are not to be trusted.

convenience, not a probability sample. The nonresponse rate is not reported. Students received course credit for participating in the survey. The sampling frame is grossly underinclusive with respect to the target population: U.S. young adults. There are less than half as many women as men; the volunteers included students as old as 36, although about 75% were freshmen; and the racial makeup of the sample does not match that of U.S. young adults (for instance, about 85% were white, compared with about 75% for the nation). There is no reason the results should be representative of other undergraduates at that institution, of students at other institutions, or of U.S. young adults.

The study acknowledges that the college students they study might not be representative of university students not in committed relationships, nor of American young adults. (Op. cit., at 449.) The study points out that dishonest reporting by students could be an issue for the accuracy of their conclusions. (Op. cit., at 448–9.)

6. Drouin, M., K.N. Vogel, A. Surbey, and J.R. Stills, 2013. Let's talk about sexting, baby: Computer-mediated sexual behaviors among young adults, *Computers in Human Behavior*, `http://dx.doi.org/10.1016/j.chb.2012.12.030`. Last visited 15 March 2013.

The sample for this study is 253 volunteers from an introductory psychology class at a midwestern university. It is a sample of convenience, not a probability sample. The response rate was not reported. The sampling frame is grossly underinclusive with respect to the target population, all U.S. young adults. The racial distribution of the sample does not match that of U.S. young adults, nor does the age distribution (75% were freshmen). There is no reason the results should be representative of other undergraduates at that institution, of students at other institutions, nor of U.S. young adults as a whole. The study notes "This sample is, by no means, representative of all American young adults." (Op. cit., at 5.)

7. Ferguson, C.J., 2011. Sexting behaviors among young Hispanic women: Incidence and association with other high-risk sexual behaviors. *Psychiatric Quarterly, 82*, 239–243. doi:10.1007/s11126-010-9165-8.

As indicated in the title, this study focuses on Hispanic women, not on young adults as a whole. The respondents were 207 Hispanic women

11

from a single university. The response rate is unknown. The sampling frame is grossly underinclusive with respect to the target population, U.S. young adults. The sample appears to be a sample of convenience, not a probability sample.

8. Gordon-Messer, D., J.A. Bauermeister, A. Grodzinski, and M. Zimmerman, 2012. Sexting among young adults, *J. Adolescent Health*, `http://dx.doi.org/10.1016/j.adohealth.2012.05.013`. Last visited 15 March 2013.

   The sample consists of roughly 2,450 volunteers recruited online through Facebook ads, using "referral chains."

   This is a sample of convenience, not a probability sample. The response rate cannot be determined, because it is not known how many people were solicited through the referral chains but did not sign up to take the survey. Dr. Zimmerman claims that since some characteristics of this sample approximately match those of other studies (that also use poor sampling methodology), that validates his sample. (Zimmerman report, at 2) Dr. Drouin makes similar claims about her studies. (Drouin report, at 5) However, matching other surveys, or even matching national demographic characteristics, is far from a guarantee that the sample even roughly matches the target population with respect to the question under study, in this case, the prevalence of sexting. (See, e.g., C. Wang, 1993. *Sense and Nonsense of Statistical Inference: Controversy, Misuse, and Subtlety*, Marcel Dekker, Inc., NY, at 172–180)

9. Lenhart, A., 2010. Teens, Adults and Sexting: Data on sending and receipt of sexually suggestive nude or nearly nude images by Americans. Pew Internet and American Life Project.
   `http://www.pewinternet.org/Presentations/2010/Oct/Teens-Adults-and-Sexting.aspx` Last visited 24 March 2013.

   The study used random-digit dialing and focus groups to reach 800 12–17 year olds, or their parents. The response rate is not reported, nor is the size of the focus groups. Random-digit dialing can be used to draw a probability sample, but the target population does not match the sampling frame in this case, and the probability that a given member of the target population is included is unknown. That makes it impossible to construct a statistically unbiased estimate of the prevalence

12

of sexting. That said, it is notable that the study reports that only 4% of teens had sent sexts, and only 15% had received a sext depicting someone they know. Note that this study focuses on minors, not young adults: The target population does not match the population of interest for this case.

This Pew presentation is based on a 2009 research report that Dr. Zimmerman does not cite:

Lenhart, A., 2009. Teens and Sexting: How and why minor teens are sending sexually suggestive nude or nearly nude images via text messaging, Pew Research Center.
`http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Teens_and_Sexting.pdf`, Last accessed 24 March 2013.

This 2009 research report indicates that the survey contacted over 64,000 households of which only 17,000 agreed to participate; of those 17,000, many were eliminated because of language barriers or because they did not have children 12–17 in the home. Of those who were not eliminated, roughly half completed the survey. Pew calculated the overall response rate to be about 11–14%: quite low. Bias is a serious concern. The cited figures for teen sending and receiving of sexts was apparently obtained from focus groups in a handful of cities—not from the probability sample, which apparently was a sample of parents of teens, not teens or young adults. There were about 600 teen cell phone users in the focus groups.

10. Mitchell, K.J., D. Finkelhor, L.M. Jones, and J. Wolak, 2012. Prevalence and characteristics of youth sexting: a national study. *Pediatrics, 129(1)*, 13–20. doi: 10.1542/peds.2011-1730.

The sample consisted of 1560 youth Internet users, age 10–17. The sample was obtained by random digit dialing. As mentioned above, random-digit dialing can be used to draw a probability sample, but the target population (young adults) does not match the sampling frame (youths whose parents were reachable by telephone and whose parents consented to the interview), and the probability that a given member of the target population is included is unknown, so it cannot be taken into account to make an unbiased estimate of the prevalence of sexting. The response rate in this survey was approximately 65%, so nonresponse bias is a serious concern. Finally, the survey consisted almost entirely

of landline users. The researchers had originally intended to sample cell phone users as well (which is increasingly important as more individuals have no landline), but they had contacted only 45 cell phone users by the time their data collection finished.

This is one of the best studies cited, although the nonresponse rate and the failure to contact cell phone users raise questions of statistical bias. Moreover, the target population is youths rather than young adults, and requiring parental consent—while required—may introduce bias. Notably, in this study, less than 10% of respondents had sent or received a sext. (Op. cit., at 4.)

11. Rice, E., H. Rhoades, H. Winetrobe, M. Sanchez, J. Montoya, A. Plant, and T. Kordic, 2012. Sexually Explicit Cell Phone Messaging associated with sexual risk among adolescents. *Pediatrics, 13(4)*, 667–673. Published online September 17, 2012. doi: 10.1542/peds.2012-0021

The participants in this study were a probability sample from Los Angeles schools, so extrapolation beyond Los Angeles schools to other schools lacks statistical foundation. The response rate to the survey was 76%, so nonresponse bias could be considerable. Results were weighted by race and ethnicity, which raises other questions. The sample was predominantly Hispanic (72%). There were fewer whites (9%) than Blacks/African-Americans (12%); this is facially unrepresentative of the nation as a whole. In the sample, approximately 14% reported having sent a sext. Again, this is a study of adolescents, not young adults: The sampling frame does not match the target population, U.S. young adults. This study cannot provide a reliable estimate of the prevalence of sexting among U.S. young adults.

12. Temple, J.R., J.A. Paul, P. Van Den Berg, V.D. Le, A. McElhany, and B.W. Temple, 2012. Teen Sexting and Its Association with Sexual Behaviors. *Archives of Pediatrics and Adolescent Medicine, 166(9)*, 828–833. Published online July 2, 2012. doi:10.1001/archpediatrics.2012.835

This seems to be a convenience sample of 948 public high school students, not a probability sample. The study does not say how the schools or students were selected, and the response rate was not reported. A disproportionate number of students were female (55.9%). The racial and ethnic characteristics were quite unlike the U.S. as a

14

whole: 26.6% Black/African-American, 30.3% white, 31.7% Hispanic, and 3.4% Asian. This is facially unrepresentative of the U.S. Moreover, this is a study of adolescents, not U.S. young adults: The sampling frame does not match the target population for this litigation, U.S. young adults.

13. The National Campaign to Prevent Teen and Unplanned Pregnancy, 2008. *Sex and Tech: Results from a survey of teens and young adults*, `http://www.thenationalcampaign.org/sextech/pdf/sextech_summary.pdf`. Last visited 23 March 2013.

   This study was also based on an online survey of volunteers—a sample of convenience with an unknown and unknowable response rate. It does not appear to have been peer-reviewed. The document itself warns:

   > "Respondents for this survey were selected from among those who have volunteered to participate in TRU's online surveys. Respondents were stratified according to the U.S. Census and the data have been weighted to reflect the demographic composition of teens and young adults. *Respondents do not constitute a probability sample.*" The National Campaign, 2008, at 5 (emphasis added).

The pattern is clear: most of the studies cited by Drs. Drouin and Zimmerman are based on samples of convenience. The studies that use probability samples—which are more reliable, despite having their own problems—generally report far lower rates of sexting than the studies based on samples of convenience. However, the studies based on probability samples had sampling frames that do not match the target population for this litigation, U.S. young adults. Therefore, extrapolation to all U.S. young adults is not justified; moreover, the response rates are so low that nonresponse bias is a serious concern for all but one of the studies. In summary, there is no reliable evidence about the rate of sexting among U.S. young adults in any of the studies Dr. Drouin or Dr. Zimmerman cites, including their own studies.

6   Discussion

To her credit, Dr. Drouin is appropriately circumspect in most of her report. For instance, she writes "[t]hese disparate results (and methods) also make

it difficult to generalize our findings to a national sample. However *if* the findings ...could be generalized ..." (Drouin report, at 4, emphasis added) Similarly, "[t]oo few research studies have examined this topic ...to make reasonable, scientific estimates of the prevalence of sexting involving visual depictions among this group." (Drouin report, at 4) "It is quite possible that the cohort of students at our undergraduate institution ...is qualitatively different somehow from young adults in other parts of the country ..." (Drouin report, at 5)

Despite these concerns, Dr. Drouin ultimately estimates that "approximately one third of *young adults* (aged 18–24) have sent text messages involving sexually-explicit visual depictions" and that she holds that opinion "to a reasonable degree of scientific certainty." (Drouin report, at 5) The word "approximately" could mean anything—a scientific estimate should have a quantified uncertainty. The phrase "reasonable degree of scientific certainty" has no scientific meaning. It is not a term of art in science or statistics.

Dr. Zimmerman extrapolates his rates of 30% and 41% to the entire U.S. young adult population, but acknowledges that "such extrapolation is sometimes tenuous." (Zimmerman report, at 2) However, he concludes that it is "very unlikely" that his estimates are off by as much as 50%. There is no scientific or statistical basis for that conclusion; even the word "likely" is not being used in a scientific sense, since that would require a probability sample, and there is no probability sample behind his estimate.

A proper scientific statement would be of the form "an unbiased estimate of the rate is $x\%$, with a margin of error of $y\%$ at 95% confidence," or "there is 95% statistical confidence that the rate is between $s\%$ and $t\%$." For the studies co-authored by Dr. Drouin and Dr. Zimmerman and most of the studies they cite, no such confidence statement can be made, because their samples are not random samples from any larger population, because the response rates are low or unknown, and because the sampling frames do not match the target population, U.S. young adults. The numbers put forward by Drs. Drouin and Zimmerman amount to guesses, not scientific estimates. They are not reliable quantitative estimates; they were not derived by following sound statistical practice; and their uncertainties cannot be quantified scientifically, for instance through a margin of error or a confidence interval (except trivially: the true rates are certainly between 0% and 100%).

Even for the studies cited that use some form of random selection, no proper statistical estimate of the rate of sexting among U.S. young adults can be made, because the samples were not drawn from the relevant population

16

(e.g., they were drawn only from a limited geographic region, or were samples of adolescents, rather than of young adults) and/or because the response rate is so low that nonresponse bias could have a large effect on the estimate. The potential biases and uncertainties are extremely large in using any of the cited studies to estimate the rate of sexting among U.S. young adults.

Dated this fourth day of April, 2013

Philip B. Stark

17